`Semantic procedure' is an oxymoron

Citation for published version:
Bundy, A 1993, '`Semantic procedure' is an oxymoron', *Behavioral and Brain Sciences*, vol. 16, no. 2. https://doi.org/10.1017/S0140525X00030338

Digital Object Identifier (DOI):
10.1017/S0140525X00030338

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published In:
Behavioral and Brain Sciences

six athletes, to check that they too turn out to be canoeists? But of course we do not – because we see that no matter how many athletes there are, each athlete has to be a baker by the first premise, and a canoeist by the second premise – but note that *that* is a logical argument which contains a variable ("each athlete") that is not allowed in J-L & B's theory. Thus, J-L & B's theory does not account for people's perception of the necessity of the conclusion without checking other models. (Of course, given some logic, no model is needed to solve that particular syllogism.)

Second, although the theory claims to invoke only specific instances, parts of the theory appear to make tacit use of variables, and even of inference rules. As one example, consider the logic of "[ ]," the "exhaustivity tag." This ensures that, given a model such as

[a] b
[a] b
. . .

the left-hand column must be fleshed out with [~a], for example,

[a]	b
[a]	b
[~a]	b
[~a]	~b

The meaning of "[ ]" appears to be given by a tacit inference rule (with variables):

$$\frac{\text{y is an unrepresented entry in a column that contains "[x]"}}{y = [\sim x]}$$

J-L & B would also need the rule:

$$[\sim\sim x] \equiv [x]$$

Thus, some mental logic is implicit in J-L & B's theory, making it a hybrid theory and blunting the logic/model opposition that J-L & B insist on.

Third, many versions of the mental-logic thesis assume that some logical apparatus is developmentally primitive – part of an innate format for representing declarative knowledge, of a syntax of thought (e.g., Braine 1990; 1992; in press; O'Brien, in press; cf. Fodor 1975). This would be consistent with the very widespread, and perhaps universal appearance in human languages of connectives similar to English *and, or, if,* and negation, in association with the same common inference forms; likewise, one tends to find words for *all* and *each*, and there are other logical elements (e.g., certain modals) that may be universal. A mental logic provides a more natural explanation than mental models of the reason why these particular elements should be so common.

Finally, J-L & B's theory requires the concept of an *unconscious* mental model held in *working memory*, a paradoxical combination: Information-processing theories customarily take the content of working memory as accessible to consciousness (e.g., Ericsson & Simon 1984). Mental models clearly are often accessible to consciousness, as any reader can attest who attempts the spatial relations problems in J-L & B's Chapter 5. (Similarly, in the folk science exemplified in Gentner & Stevens [1983], the models were generally accessible to consciousness.) However, for logic problems, introspection suggests that subjects do not consistently use the kinds of models proposed. Models are usually not reported in propositional problems (in our experience – cf. Braine et al. 1984); for other logic problems, old evidence (Storring, cited in Woodworth 1938), which is consistent with unpublished work of O'Brien and myself, suggests that there is great variation across people and problems in whether models are reported, and in the kind of model reported. I cannot help thinking that the purpose of the unconsciousness postulate is to shield the theory from this obvious kind of evidence. At the very least, J-L & B need to explain the variability in what people report.

In sum, I have argued, first, that mental models cannot suffice for reasoning, given J-L & B's notion of model; second, that some mental logic is implicit in J-L & B's theory, making it more of a hybrid than they allow; third, that the common logical apparatus of human languages argues for a mental logic and is hard to explain using J-L & B's theory; and, finally, that there is a deep problem with J-L & B's rejection of introspective evidence.

"Semantic procedure" is an oxymoron

Alan Bundy
Department of Artificial Intelligence, University of Edinburgh, Edinburgh EH1 1HN, Scotland
Electronic mail: *bundy@ed.ac.uk*

1. Introduction. Johnson-Laird & Byrne (J-L & B) are to be congratulated on proposing a new mechanism for deductive inference and for presenting extensive evidence for the psychological validity of this mechanism. I will have no quarrel with this mechanism or with the psychological claims; both deserve attention and further investigation. My argument is against the implied epistemic nature of the new mechanism.

In *Deduction*, the mental-model mechanism is described as a "semantic procedure" (p. 23) and is said to be "compatible with the way in which logicians formulate a semantics for a calculus" (p. 36). Mental models are contrasted strongly with rule-based mechanisms (e.g., pp. 23, 195). The implication, whether intended or not, is that the mental-model mechanism directly addresses the problem of intentionality. A mental-model-based computer program, it seems, would automatically give meaning to computational states.

I will argue that this implication is wrong; mental models have no more to say about intentionality than rule-based mechanisms. The attachment of the adjective "semantic" to a deductive mechanism, or to any computer program, is misleading and confusing. The phrase "semantic procedure" is an oxymoron. Mental-model and rule-based mechanisms differ only in degree and not in kind.

2. The meanings of "semantics." Unfortunately, the issue is clouded because the word "semantics" is used in different ways by different communities. For example, logicians use it to describe a mapping from the expressions of a logical theory to the "meaning" of these expressions. To give a semantics to a logic is to provide this mapping. Tarski provided a semantics for predicate calculus by showing how logical sentences in a theory could be mapped to truth or falsity in a model.

There is an ambiguity about whether these models are aspects of the real world or mathematical theories in their own right. For a semantics to map formulae to their meaning, models should be part of the real world. However, there are several forces encouraging their formalisation as mathematical theories. Formulae in commonsense reasoning are relatively easy to map to the real world. For example, in *loves (John, Mary)* the constants *John* and *Mary* map to specific individuals John and Mary, *loves* to the relationship of loving and *loves (john, mary)* to the assertion that John loves Mary. Mathematical formulae, for example, 2 + 2 = 4, are harder to map to the real world because the coherence of the mapping presupposes a platonic commitment to the existence of 2, 4, and so on. Couple this with the natural tendency of mathematicians to formalise, and it becomes easier for them to regard models as mathematical theories of sets of objects on which functions and relations are defined. The sense of "semantics" in which it assigns "meaning" is then lost.

Linguists generally use "semantics" to describe, not the mapping to a meaning, but the meaning itself. A semantic representation of a natural language sentence is contrasted with the syntactic representation. The syntactic representation is the original string of words or a parse tree with these words labelling the leaves. The semantic representation must capture not this grammatical structure but its content. Confusingly, this is usually done by a logical formula; so the linguist's semantics is the logician's syntax!

Computer scientists use the word "semantics" to describe the mapping from a programming language to a mathematical theory. Ironically, this turns the logician's usage on its head. Logical semantics translates a mathematical formula into a program for calculating a truth value; computer science semantics translates a program into a mathematical formula.

Because of their remark on p. 36 of *Deduction* (see para. 1 above), I will assume that J-L & B intend the word "semantics" in the logician's sense. I assume that their mental models are based on Tarski's models of logical theories; that their deductive mechanism is an attempt to reason in the model theory in contrast to rule-based mechanisms that reason in the proof theory. I claim that it is not possible to do this.

3. Is semantic reasoning possible? If we regard Tarskian models as part of the real world, then reasoning with them would entail physically manipulating the real world. This has limited utility. It is not possible to conduct forward planning, hypothetical reasoning, counterfactual reasoning or abstract reasoning by manipulating the current world state. We must reason by manipulating an internal representation of the world.

At this point the problems of intentionality emerge, that is, we need a semantics to map this internal representation onto its meaning. This remains true even if the internal representation is based on a Tarskian model. Calling the manipulation procedure "semantic" does not affect the situation.

Basing a computational reasoning mechanism on Tarskian models presents problems for a finite computer. For example, some models have an infinite domain of objects. Some reasoning involves proving that an infinite collection of objects has a property. Some reasoning involves the representation and use of incomplete or vague information. These problems are solved in rule-based mechanisms by the use of quantifiers, variables, disjunction, and so forth. Some equivalent device is needed in model-based reasoners if they are to have the same reasoning power. J-L & B use such devices in their mental-model mechanism. For example, infinite numbers of objects are represented by a finite number of tokens; incomplete information is represented by having alternative models to cover the range of possibilities.

4. Are rule- and model-based reasoners different in kind? One paradigmatic example of a rule-based deductive system is a resolution-based theorem prover. The rules are formulae of predicate calculus in clausal form representing the axioms of the theory and the negation of the conjecture. The conjecture is proved by *reductio ad absurdum;* the clauses are "resolved" together, usually exhaustively, until the empty clause is derived.

However, resolution can also be viewed as a systematic attempt to check that none of the models of the theory provide a counterexample to the conjecture. The fact that resolution can be viewed in this way goes back to a metalogical theorem of Herbrand's. If the attempt to prove the conjecture fails after a finite search then a counterexample to the conjecture can be read off automatically from the failed attempt. Thus resolution can be viewed both as a rule-based and as a model-based mechanism!

This potential duality was brought home to me forcibly as a result of my first foray into automatic theorem proving. I built a model-based theorem prover for arithmetic called SUMS (Bundy 1973). Its model consisted of a representation of the "real line" as used by mathematicians in informal blackboard arguments. The hypotheses of the theorem were represented by placing points in appropriate positions on this "real line" and the conclusion was then read off from the model.

As I tried to get SUMS to prove harder and harder theorems, this simple idea became more and more elaborate. For example, consequences of the original hypotheses had to be propagated around the model before the conclusion could be read off. The natural propagation mechanism was forward-chaining with rules. After a while I realised that I had just built yet another rule-based mechanism. SUMS was now similar to a standard semantic tableau prover with a bottom-up search strategy. SUMS' progression from model-based to rule-based was incremental. There was no point at which the nature of its reasoning dramatically changed in kind.

5. Conclusion. I have argued that there is no difference in kind between the mental-model deduction mechanism of J-L & B and rule-based mechanisms. Indeed, it is possible to view many deduction mechanisms as simultaneously of both types. The issue of intentionality arises with both types of mechanism, and is not finessed by the use of a model-based approach. To the best of my knowledge J-L & B make no claim to the contrary. However, others may erroneously draw that conclusion from the free use of words like "semantics," "model," and so forth. For this reason I recommend that the word "semantics" be used with extreme caution. It is a highly ambiguous term and has great potential to mislead.

None of this detracts from Johnson-Laird & Byrne's significant contribution in defining a new deduction mechanism and providing evidence for its psychological validity.

ACKNOWLEDGMENT

The work reported in this commentary was supported by a SERC Senior Fellowship to the author.

Mental models and nonmonotonic reasoning

Nick Chater

Department of Psychology, University of Edinburgh, Edinburgh EH8 9JZ, Scotland

Electronic mail: *nicholas@cogsci.ed.ac.uk*

Johnson-Laird & Byrne (J-L & B) are equivocal concerning the scope of mental-model theory. On the one hand, they are careful to note that mental models are aimed primarily at explaining deduction, although commonsense inference is not deductive in character. On the other hand, they contend that mental-model theory solves the problem of nonmonotonic reasoning, which is *not* deductive and is characteristic of commonsense inference. This equivocation requires clarification: An account of deductive reasoning casts light on a fascinating if rather arcane human ability; an account of nonmonotonic inference in general would be little short of a theory of thinking. It is not clear, therefore, exactly what J-L & B see as the domain of the mental-model account. I shall argue that mental-model theory does not in fact address the problem of understanding commonsense nonmonotonic reasoning, still less provide a solution to it.

Everyday, commonsense reasoning may be conceived of as a species of inference to the best explanation: It involves inferring from given information to what best explains and is best explained by that information (Fodor 1983). Such inference is nonmonotonic, because the addition of new information can invalidate what were previously plausible conclusions. So, for example, the plausible inference from hearing the sound of purring behind the door to the conclusion that the cat is trapped in the cellar is immediately overridden if I catch sight of the cat in the garden. The premise on which the inference is based, the purring, need not be withdrawn, although another explanation for this fact may be sought. By contrast, in monotonic reasoning, the conclusion of a valid argument can only be challenged if one of its premises is false.

Providing an account of inference to the best explanation is very difficult. Inference to the best explanation encompasses